January 6, 1990 LIDS-P-1938

Successive Projection under a Quasi-Cyclic Order *

by

Paul Tseng**

Abstract

A classical method for finding a point in the intersection of a finite collection of closed convex sets is the successive projection method. It is well-known that this method is convergent if each convex sets is chosen for projection in a cyclical manner. In this note we show that this method is still convergent if the length of the cycle grows <u>without bound</u>, provided that the growth is not too fast. Our argument is based on an interesting application of the Cauchy-Schwartz inequality.

KEY WORDS: successive projection, convex set, quasi-cyclic order.

*This research is supported by the U.S. Army Research Office, contract DAAL03-86-K-0171 (Center for Intelligent Control Systems).

**Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **06 JAN 1990**	2. REPORT TYPE	3. DATES COVERED **00-01-1990 to 00-01-1990**
4. TITLE AND SUBTITLE **Successive Projection under a Quasi-Cyclic Order**		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Massachusetts Institute of Technology,Laboratory for Information and Decision Systems,77 Massachusetts Avenue,Cambridge,MA,02139-4307**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**		
13. SUPPLEMENTARY NOTES		
14. ABSTRACT		
15. SUBJECT TERMS		

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**11**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

1. Introduction

A fundamental problem in convex programming is that of finding a point in the intersection of a finite collection of closed convex sets. This problem has many application areas, including image reconstruction [SSW77], [Her80], linear prediction theory, multigrid methods, computed tomograph [Deu85], optimal control [GPR67], machine learning [MiP88], and linear/quadratic programming [Bre65]. One classical method for solving this problem is the successive projection method, whereby an arbitrary starting point is successively projected onto the individual convex sets to generate a sequence of points converging to a solution. This method was first proposed by Kacmarz [Kac37] for the special case where the sets are linear varieties (i.e., translates of subspaces), and was rediscovered by von Neumann [voN50], by Agmon [Agm54], and by Motzkin and Schoenberg [MoS54]. (Also see [Deu85], [Gof80], [Gof82], [Hal62], [Man84], [Mer62], [SSW77], [Tan71] for more detailed treatments of the linear case, including rate of convergence analysis [Agm54], [Deu85], [Gof80], [Gof82], [Man84], [Mer62], [SSW77] and finite convergence analysis [MoS54], [Gof80], [Gof82].) Extensions of this method to arbitrary convex sets are discussed in [Bre65], [Ere65], [Ere66], [GeS66], [GPR67], [Pol69], amongst which the analysis given in [GPR67] is the most extensive. This method can also be applied to problems in a product space to obtain a highly parallelizable method of barycentres [Pie84].

In all the existing successive projection methods, the sets are chosen for projection either in an essentially cyclic order (i.e., every set is chosen at least once every B iterations, for some fixed B > 0) or according to a maximal distance rule (i.e., choose a set that is in some sense farthest away from the current point). (Numerical evidence suggests that the essentially cyclic order is perhaps the more efficient [Man84].) In this note we show that the essentially cyclic order can be further extended to one that allows the length of each cycle, namely B, to increase without bound, provided that the rate of increase is not too fast. This extension, apart from its theoretical appeal, has the practical advantage that it allows sets for which projection is expensive to be left out of the computation increasingly more often.

2. Algorithm Description

Let $\mathcal{H}$ be a real Hilbert space endowed with an inner product $\langle\cdot,\cdot\rangle$ and let $\|\cdot\|$ denote the norm induced by the inner product (i.e., $\|x\| = \sqrt{\langle x,x\rangle}$). Let $C_1, C_2, \ldots, C_m$ be a given collection of closed convex sets in $\mathcal{H}$. Our problem is to find a point in $C = C_1 \cap C_2 \cap \ldots \cap C_m$. We make the following standing assumption regarding the C_i's.

Assumption A (feasibility): $C_1 \cap C_2 \cap \ldots \cap C_m \neq \emptyset$.

Assumption A is fairly standard. Some results for the case where the C_i's do not intersect are given in [GPR67].

We describe the successive projection method below. In this method, we begin with an arbitrary x(0) in $\mathcal{H}$ and we generate a sequence of points {x(0), x(1), ...} in $\mathcal{H}$ according to the iteration:

$$y(t) = \operatorname{argmin}\{\ \|y - x(t)\| \mid y \in C_{\sigma(t)}\ \}, \tag{1a}$$

$$x(t+1) = \omega(t)\, y(t) + (1-\omega(t))\, x(t), \tag{1b}$$

where $\sigma(t)$ is an element of $\{1,2,\ldots,m\}$ ($\sigma(t)$ specifies the set onto which projection is made at the t-th iteration), and $\omega(t)$ is a scalar (called a <u>relaxation</u> parameter) satisfying

$$\varepsilon \le \omega(t) \le 2-\varepsilon, \quad \forall\, t, \tag{1c}$$

for some fixed $\varepsilon \in (0,1]$. The relaxation mechanism $\omega(t)$ was first introduced in [Agm54] and in [MoS54]. It has been observed that, in certain cases, a value of $\omega(t)$ different from one (i.e., <u>under/over-relaxation</u>) can significantly improve the convergence [Gof80], [Her80], [Man84].

The iterates $\{x(t)\}$ in general do not converge unless we impose certain restrictions on the order in which projection onto the sets $C_1, C_2, \ldots, C_m$ are made. We will consider the following order of projection, introduced in [TsB87]:

Quasi-Cyclic Order: There exists a sequence of integers $\{\tau_1, \tau_2, \ldots\}$ satisfying

$$\tau_1 = 0, \qquad \tau_{k+1} - \tau_k \geq m \;\; \forall\, k, \qquad \sum_{k=1}^{\infty} 1/(\tau_{k+1} - \tau_k) = \infty, \tag{2a}$$

for which

$$\{1,2, \ldots,m\} \subseteq \{\sigma(\tau_k), \sigma(\tau_k+1), \ldots, \sigma(\tau_{k+1}-1)\}, \quad \forall\, k. \tag{2b}$$

Roughly speaking, the quasi-cyclic order (of projection) states that every set C_i must be chosen for projection at least once between the τ_k-th and the $(\tau_{k+1}-1)$-st iteration (called the k-th <u>quasi-cycle</u>) for all k [cf. (2b)], and that the length of the k-th quasi-cycle, namely $\tau_{k+1} - \tau_k$, cannot grow too fast with k [cf. (2a)]. One particular choice of the τ_k's, namely $\tau_k = m(k-1)$ for all k, gives rise to the well-known <u>cyclic</u> order (of projection), for which $\sigma(t) = t \pmod m + 1$ for all t (and the length of each quasi-cycle is exactly m). A more interesting choice of the τ_k's is given by

$$\tau_{k+1} = \tau_k + k\,m, \qquad \forall\, k,$$

for which the length of the k-th quasi-cycle <u>increases linearly</u> with k.

3. Convergence Analysis

Below we give the main result of this note.

Theorem 1. Let $\{x(t)\}$ be a sequence of iterates generated by (1a)-(1c) using the quasi-cyclic order of projecton. Then, $\{x(t)\}$ converges in the weak topology to a point x^∞ in C.

Proof: Fix any $\bar{x} \in C$. For any nonnegative integer t, we have from (1a) that y(t) is the orthogonal projection of x(t) onto $C_{\sigma(t)}$. Since $\bar{x} \in C$ so that $\bar{x} \in C_{\sigma(t)}$ (cf. definition of C), this implies

$$\langle y(t) - x(t), \bar{x} - y(t) \rangle \geq 0.$$

Since [cf. (1b)]

$$\|\bar{x} - x(t)\|^2 = \|\bar{x} - x(t+1)\|^2 + 2\omega(t) \langle y(t) - x(t), \bar{x} - y(t) \rangle + \omega(t)(2-\omega(t)) \|y(t) - x(t)\|^2,$$

this, together with (1c), implies

$$\|\bar{x} - x(t)\|^2 \geq \|\bar{x} - x(t+1)\|^2 + \varepsilon^2 \|x(t+1) - x(t)\|^2, \qquad \forall\, t \geq 0, \tag{3}$$

so that $\{x(t)\}$ is bounded, $\|\bar{x} - x(t)\|$ is nonincreasing with t, and

$$\sum_{t=0}^{\infty} \|x(t+1) - x(t)\|^2 < \infty. \tag{4}$$

We claim that there exists a subsequence $\mathcal{K}$ of $\{1,2,\ldots\}$ for which

$$\sum_{t=\tau_k+1}^{\tau_{k+1}} \|x(t+1) - x(t)\| \to 0 \quad \text{as } k \to \infty,\ k \in \mathcal{K}. \tag{5}$$

To see this, suppose that such a subsequence does not exist. Then, there would exist a positive scalar δ and an integer $\bar{k}$ such that

$$\delta \le \sum_{t=\tau_k+1}^{\tau_{k+1}} \|x(t+1) - x(t)\|, \qquad \forall\, k \ge \bar{k}.$$

Since, by the Cauchy-Schwartz inequality, there holds

$$\sum_{t=\tau_k+1}^{\tau_{k+1}} \|x(t+1) - x(t)\| \le \sqrt{\sum_{t=\tau_k+1}^{\tau_{k+1}} \|x(t+1) - x(t)\|^2} \cdot \sqrt{\tau_{k+1} - \tau_k}\,,$$

this implies

$$\delta^2 \le \sum_{t=\tau_k+1}^{\tau_{k+1}} \|x(t+1) - x(t)\|^2 \ (\tau_{k+1} - \tau_k), \qquad \forall\, k \ge \bar{k},$$

so that

$$\delta^2 \sum_{k=\bar{k}}^{\infty} 1/(\tau_{k+1} - \tau_k) \le \sum_{k=\bar{k}}^{\infty} \Big[\sum_{t=\tau_k+1}^{\tau_{k+1}} \|x(t+1) - x(t)\|^2 \Big]$$

$$= \sum_{t=\tau_{\bar{k}}+1}^{\infty} \|x(t+1) - x(t)\|^2. \tag{6}$$

The left hand side of (6), by (2a), has the extended value of ∞, while the right hand side of (6) according to (4) has finite value, thereby reaching a contradiction. Hence, (5) holds for some $\mathcal{K} \subseteq \{1,2,\ldots\}$.

Let $\mathcal{K}$ be any subsequence of $\{1,2,\ldots\}$ satisfying (5). Since $\{x(t)\}$ is bounded [cf. (3)], there exist some $x^\infty \in \mathcal{H}$ and some $\mathcal{K}' \subseteq \mathcal{K}$ such that

$$\{x(\tau_k+1)\}_{k\in\mathcal{K}'} \text{ converges weakly to } x^\infty. \tag{7}$$

We claim that $x^\infty \in C$. To see this, fix any $i \in \{1,2,\ldots,m\}$. Since the projections are in the quasi-cyclic order, then for each integer $k \ge 1$ there exists some $\rho_k \in \{\tau_k, \tau_k+1, \ldots, \tau_{k+1}-1\}$ satisfying

$\sigma(\rho_k) = i$ [cf. (2b)]. By using the triangle inequality together with the fact $\|x(\rho_k+1) - y(\rho_k)\| \le (1/\varepsilon-1)\|x(\rho_k+1) - x(\rho_k)\|$ for all $k \ge 1$ [cf. (1b), (1c)], we have

$$\begin{aligned} \|x(\tau_k+1) - y(\rho_k)\| &\le \sum_{t=\tau_k+1}^{\tau_{k+1}} \|x(t+1) - x(t)\| + \|x(\rho_k+1) - y(\rho_k)\| \\ &\le \sum_{t=\tau_k+1}^{\tau_{k+1}} \|x(t+1) - x(t)\| + (1/\varepsilon-1)\|x(\rho_k+1) - x(\rho_k)\|, \qquad \forall\, k \ge 1. \quad (8) \end{aligned}$$

Eqs. (5), (8) together with $\{\|x(t+1) - x(t)\|\} \to 0$ [cf. (4)] imply

$$\lim_{k\to\infty, k\in\mathcal{K}'} \|x(\tau_k+1) - y(\rho_k)\| = 0,$$

which, combined with (7), yields

$$\lim_{k\to\infty, k\in\mathcal{K}'} \langle u, y(\rho_k)\rangle = \lim_{k\to\infty, k\in\mathcal{K}'} \langle u, x(\tau_k+1)\rangle = \langle u, x^\infty\rangle, \qquad \forall\, u \in \mathcal{H},$$

so that $\{y(\rho_k)\}_{k\in\mathcal{K}'}$ converges weakly to x^∞. Since $y(\rho_k) \in C_i$ [cf. $\sigma(\rho_k) = i$ and (1a)] for all k and C_i is closed and convex, this shows $x^\infty \in C_i$ [DuS66, p. 422, Theorem 3.13]. Since the choice of i was arbitrary, we obtain $x^\infty \in C_i$ for all i, and therefore $x^\infty \in C$.

We now show that $\{x(t)\}$ has a unique weak limit point. Our argument follows that given in [Bre65] and is presented here for completeness. Suppose that $\{x(t)\}$ does not have a unique limit point. Then, there exist $x_1^\infty \in C$, $x_2^\infty \in C$ with $x_1^\infty \ne x_2^\infty$ and subsequences $\{x(t)\}_{t\in\mathcal{T}_1}$, $\{x(t)\}_{t\in\mathcal{T}_2}$ converging weakly to, respectively, x_1^∞ and x_2^∞. By replacing $\bar{x}$ in (3) by x_1^∞, we find that $\|x_1^\infty - x(t)\|$ is nonincreasing with t so that there exists a scalar α_1 such that

$$\{\|x_1^\infty - x(t)\|^2\} \to \alpha_1. \qquad (9a)$$

Similarly, by replacing $\bar{x}$ in (3) by x_2^∞, we find that there exists a scalar α_2 such that

$$\{\|x_2^\infty - x(t)\|^2\} \to \alpha_2. \qquad (9b)$$

Now, for any $t \in \mathbf{T}_1$ we have

$$\|x_2^\infty - x(t)\|^2 \quad = \quad \|x_2^\infty - x_1^\infty\|^2 + 2\langle x_2^\infty - x_1^\infty , x_1^\infty - x(t)\rangle + \|x_1^\infty - x(t)\|^2,$$

so that, by letting $t \to \infty$, $t \in \mathbf{T}_1$, we obtain from (9a)-(9b) and the weak convergence of $\{x(t)\}_{t \in \mathbf{T}_1}$ to x_1^∞ that $\alpha_2 = \|x_2^\infty - x_1^\infty\|^2 + \alpha_1$. By an analogous argument with the role of x_1^∞ and x_2^∞ reversed, we also obtain $\alpha_1 = \|x_1^\infty - x_2^\infty\|^2 + \alpha_2$. Adding these two relations yields $0 = \|x_1^\infty - x_2^\infty\|^2$ and hence $x_1^\infty = x_2^\infty$, a contradiction. Q.E.D.

We remark that <u>strong</u> convergence of the sequence $\{x(t)\}$ generated by (1a)-(1c), using the quasi-cyclic order of projection, can also be established under the same set of conditions on the C_i's as those given by Gubin et. al. [GPR67, Theorem 1]. The proof of this follows (in a straightforward manner) from combining the proof of Theorem 1 with that of [GPR67, Theorem 1].

References

[Agm54] Agmon, S., "The Relaxation Method for Linear Inequalities," *Canadian J. Math.*, 6 (1954), 382-392.

[Bre65] Bregman, L. M., " The Method of Successive Projection for Finding a Common Point of Convex Sets", *Akad. Nauk SSSR Dokl.*, 162 (1965), 487-490; *Soviet Math. Doklady*, 162 (1965), 688-692.

[Deu85] Deutsch, F., "Rate of Convergence of the Method of Alternating Projections," in *Parametric Optimization and Approximation* (B. Brosowski and F. Deutsch, eds.), Birkhauser Verlag, Boston (1985).

[DuS66] Dunford, N. and Schwartz, J. T., *Linear Operators I: General Theory* (third edition), Wiley-Interscience, New York, NY (1966).

[Ere65] Eremin, I. I., "A Generalization of the Motzkin-Agmon Relaxational Method," *Uspekhi. Mat. Nauk.*, 20 (1965), 183-187 (in Russian).

[Ere66] Eremin, I. I., "On Systems of Inequalities with Convex Functions in the Left Sides," *Amer. Math. Soc. Transl.*, 88 (1966), 67-83. Translation of *Izv. Akad. Nauk SSSR Ser. Mat.*, 30 (1966), 265-278.

[GeS66] Germanov, M. A. and Spiridonov, V. S., "On a Method of Solving Systems of Non-Linear Inequalities," *Z. Vycisl. Mat. i. Mat. Fiz.*, 6 (1966), 335-336; *USSR Comp. Math. and Math. Phys.*, 2 (1966), 194-196.

[Gof80] Goffin, J. L., "The Relaxation Method for Solving Systems of Linear Inequalities," *Math. Oper. Res.*, 5 (1980), 388-414.

[Gof82] Goffin, J. L., "On the Nonpolynomiality of the Relaxation Method for Systems of Linear Inequalities," *Math. Prog.*, 22 (1982), 93-103.

[GPR67] Gubin, L. G., Polyak, B. T. and Raik, E. V., "The Method of Projections for Finding the Common Point of Convex Sets," *Z. Vycisl. Mat. i Mat. Fiz.*, 7 (1967), 1211-1228; *USSR Comp. Math. and Math. Phys.*, 6 (1967), 1-24.

[Hal62] Halperin, I., "The Product of Projection Operators," *Acta Sci. Math.* (Szeged), 23 (1962), 96-99.

[Her80] Herman, G. T., *Image Reconstructions from Projections, the Fundamentals of Computerized Tomography*, Academic Press, New York, NY (1980).

[Kac37] Kacmarz, S., *Bull. intern. Acad. Polonaise des Sciences A* (1937), 355-357.

[Man84] Mandel, J., "Convergence of the Cyclical Relaxation Method for Linear Inequalities," *Math. Prog.*, 30 (1984), 218-228.

[Mer62] Merzlyakov, Yu. I., "On a Relaxation Method of Solving Systems of Linear Inequalities," *Z. Vycisl. Mat. i Mat. Fiz.*, 2 (1962), 482-487; *USSR Comp. Math. and Math. Phys.*, 2 (1963), 504-510.

[MiP88] Minsky, M. L. and Papert, S. A., *Perceptrons: An Introduction to Computational Geometry* (expanded edition), The MIT Press, Cambridge, MA (1988).

[MoS54] Motzkin, T. S. and Schoenberg, I. J., "The Relaxation Method for Linear Inequalities," *Canadian J. Math.*, 6 (1954), 393-404.

[Pie84] Pierra, G., "Decomposition Through Formalization in a Product Space," *Math. Prog.*, 28 (1984), 96-115.

[Pol69] Polyak, B. T., "Minimization of Unsmooth Functionals," *Z. Vycisl. Mat. i Mat. Fiz.*, 9 (1969), 509-521; *USSR Comp. Math. and Math. Phys.*, 9 (1969), 14-29.

[SSW77] Smith, K. T., Solmon, D. C., and Wagner, S. L., "Practical and Mathematical Aspects of the Problem of Reconstructing Objects from Radiographs," *Bull. Amer. Math. Soc.*, 83 (1977), 1227-1270.

[Tan71] Tanabe, K., "Projection Method for Solving a Singular System of Linear Equations and Its Applications," *Numer. Math.*, 17 (1971), 203-214.

[TsB87] Tseng, P. and Bertsekas, D. P., "Relaxation Methods for Problems with Strictly Convex Separable Costs and Linear Constraints," *Math. Prog.*, 38 (1987), 303-321.

[voN50] von Neumann, J., *Functional Operators -- Vol. II. The Geometry of Orthogonal Spaces*, Annals of Math. Studies #22, Princeton University Press (1950). (This is a reprint of mimeographed lecture notes first distributed in 1933.)